

# Patent Abstracts of Japan

PUBLICATION NUMBER : 2000165796  
PUBLICATION DATE : 16-06-00

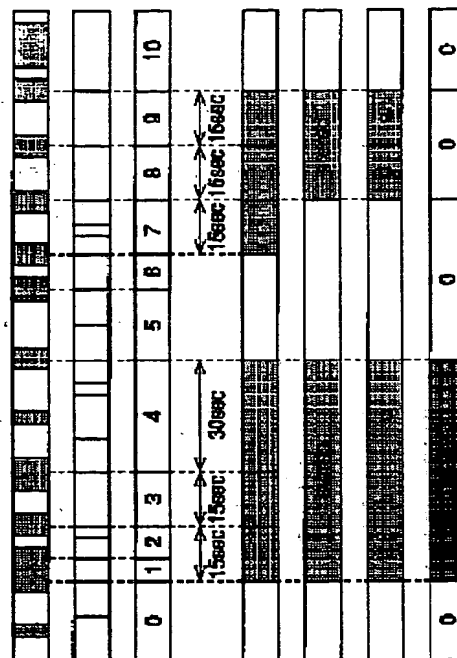
APPLICATION DATE : 30-11-98  
APPLICATION NUMBER : 10339273

APPLICANT : SONY CORP;

INVENTOR : MIYATA MASANARI:

INT.CL. : H04N 5/7826

TITLE : INFORMATION PROCESSOR AND  
INFORMATION PROCESSING  
METHOD AND PROVIDING MEDIUM



(A) (B) (C) (D) (E) (F) (G)

**ABSTRACT :** PROBLEM TO BE SOLVED: To surely detect CM included in television broadcasting.

**SOLUTION:** A CM candidate block detecting part extracts silent blocks as shown in (A), and extracts scene change points as shown in (B), and divides the scene change blocks for each scene change point in the silent blocks. Also, the CM candidate block detecting part judges the blocks grouped by the integral times as long as 15 seconds and constituted as stereo blocks as CM candidate blocks as shown in (F). Moreover, the CM candidate block detecting part compares the number of times of the scene change in the CM candidate blocks (CM candidate blocks (a), (b), (d), and (e)) at the both edges of the adjacent CM candidate blocks with a prescribed threshold (for example, 1), and judges the CM candidate blocks (CM candidate blocks (a) and (c)) in which the number of times of scene change is more than the threshold as CM blocks as shown in (G). Also, the candidate block(b) interposed between the CM candidate blocks (a) and (c) is also judged as the CM block.

COPYRIGHT: (C)2000,JPO

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2000-165796

(P2000-165796A)

(43)公開日 平成12年6月16日(2000.6.16)

(51)Int.Cl.<sup>7</sup>

H 0 4 N 5/7826

識別記号

F I

H 0 4 N 5/782

テマコード\*(参考)

A 5 C 0 1 8

Z

審査請求 未請求 請求項の数4 O L (全 16 頁)

(21)出願番号 特願平10-339273

(22)出願日 平成10年11月30日(1998.11.30)

(71)出願人 000002185

ソニー株式会社

東京都品川区北品川6丁目7番35号

(72)発明者 太田 正志

東京都品川区北品川6丁目7番35号 ソニー株式会社内

(72)発明者 水藤 太郎

東京都品川区北品川6丁目7番35号 ソニー株式会社内

(74)代理人 100082131

弁理士 稲本 義雄

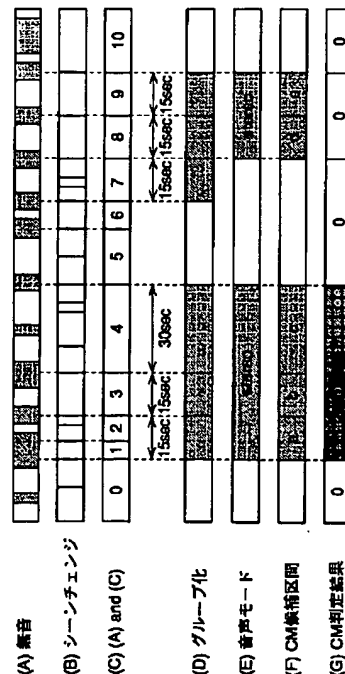
最終頁に続く

(54)【発明の名称】 情報処理装置および方法、並びに提供媒体

(57)【要約】

【課題】 テレビジョン放送に含まれるCMを確実に検出する。

【解決手段】 CM候補区間検出部は、図14(A)に示すように、無音区間を抽出し、図14(B)に示すように、シーンチェンジ点を抽出して、図14(C)に示すように、無音区間におけるシーンチェンジ点毎にシーンチェンジ区間を区切る。また、CM候補区間検出部は、図14(F)に示すように、15秒の整数倍にグループ化され、且つ、ステレオ区間である区間をCM候補区間と判定する。さらに、CM候補区間検出部は、図14(G)に示すように、隣接するCM候補区間の両端のCM候補区間(いまの場合、CM候補区間a、c、d、e)におけるシーンチェンジの回数を所定の閾値(例えば、1)と比較して、シーンチェンジの回数が閾値よりも大きいCM候補区間(いまの場合、CM候補区間a、c)をCM区間と判定する。なお、CM候補区間a、cに挟まれているCM候補区間bもCM区間と判定される。



## 【特許請求の範囲】

【請求項1】 テレビジョン放送に含まれるコマーシャルを検出する情報処理装置において、前記テレビジョン放送に関係する信号を用いて閾値を決定する閾値決定手段と、前記閾値を用いて前記テレビジョン放送の無音区間を判定する判定手段と、前記テレビジョン放送のシーンチェンジを検出するシーンチェンジ検出手段と、前記シーンチェンジ検出手段の検出結果、および前記判定手段の判定結果に基づいてコマーシャルを検出するコマーシャル検出手段とを備えることを特徴とする情報処理装置。

【請求項2】 所定の区間の前記シーンチェンジの回数をカウントするカウント手段をさらに備え、前記コマーシャル検出手段は、前記シーンチェンジの回数に基づいて前記コマーシャルを検出することを特徴とする請求項1に記載の情報処理装置。

【請求項3】 テレビジョン放送に含まれるコマーシャルを検出する情報処理装置の情報処理方法において、前記テレビジョン放送に関係する信号を用いて閾値を決定する閾値決定ステップと、前記閾値を用いて前記テレビジョン放送の無音区間を判定する判定ステップと、前記テレビジョン放送のシーンチェンジを検出するシーンチェンジ検出ステップと、前記シーンチェンジ検出ステップの検出結果、および前記判定ステップの判定結果に基づいてコマーシャルを検出するコマーシャル検出ステップとを含むことを特徴とする情報処理方法。

【請求項4】 テレビジョン放送に含まれるコマーシャルを検出する情報処理装置に、前記テレビジョン放送に係る信号を用いて閾値を決定する閾値決定ステップと、前記閾値を用いて前記テレビジョン放送の無音区間を判定する判定ステップと、前記テレビジョン放送のシーンチェンジを検出するシーンチェンジ検出ステップと、前記シーンチェンジ検出ステップの検出結果、および前記判定ステップの判定結果に基づいてコマーシャルを検出するコマーシャル検出ステップとを含む処理を実行させるコンピュータが読み取り可能なプログラムを提供することを特徴とする提供媒体。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、情報処理装置および方法、並びに提供媒体に関し、特に、テレビジョン放送に含まれるコマーシャルを検出する情報処理装置および方法、並びに提供媒体に関する。

## 【0002】

【従来の技術】例えば、録画されたテレビジョン放送を再生するとき、本編（番組）だけを連続して見たいというユーザのニーズに対応して、コマーシャル（以下、CMと記述する）を早送りする、いわゆるCMカット機能付きVCR（Video Cassette Recorder）が存在する。

【0003】そのようなVCRに用いられているCM検出アルゴリズムは、次に示すような、絶対的な、または大多数のCMが有する特徴に基づくものである。すなわち、各CMの開始時および終了時には0.1乃至2.0秒程度の無音区間が存在する、無音区間には画像のシーンチェンジ点が存在する、各CMの所要時間は15秒の整数倍である、CMは音声多重モードがステレオモードである、等の特徴が確認された場合、その部分がCMとして検出される。

【0004】また、欧米におけるテレビジョン放送では、本編とCMの間に、ブラックフレームまたはブルーフレームが存在するという絶対的な条件に基づいてCMが検出されている。

## 【0005】

【発明が解決しようとする課題】したがって、従来のCM検出アルゴリズムによれば、例えば、音声多重モードがモノラルであるCMのように、上述した特徴が含まれていないCMは検出できない課題があった。

【0006】また、従来のCM検出アルゴリズムによれば、本編の中に上述した特徴が含まれていた場合、その部分をCMとして検出してしまふ課題があった。

【0007】さらに、例えば、CMの所要時間が14秒の整数倍となったり、欧米においてブラックフレームやブルーフレームの挿入が廃止される等、従来のアルゴリズムにおいて用いていたCMの特徴が変更、または廃止された場合、CMを検出することができなくなる課題があった。

【0008】また、従来のCM検出アルゴリズムにおける無音区間の検出方法は、ある一定区間の音声の平均レベルを求め、その平均レベルが所定の閾値以下である場合、その区間を無音区間として検出していた。このため、テレビジョン放送の受信状況が悪いとき、すなわち、弱電界で音声信号のS/Nが悪い状況と、弱電界ではないときで無音声区間の検出精度が異なるので、結果的に正確にCMを検出できない課題があった。

【0009】また、従来のCM検出アルゴリズムにおけるCMの所要時間の計測は、1秒当たり約30フレームとして、フレーム数をカウントすることにより行われているが、実際の放送においては、数フレーム程度の誤差が生じるので、所要時間を判定する閾値には誤差に対応するあそびが設けられている。したがって、そのあそびに起因して、CMの誤検出、または未検出が発生する課題があった。

【0010】さらに、CMと本編の間に5秒程度のテレビ局の宣伝用スポットが放送された場合、そのスポット

は視聴者にとってはCMと同様なものであるが、上述したCMの特徴を有していないので検出できない課題があった。

【0011】本発明はこのような状況に鑑みてなされたものであり、テレビジョン放送に含まれるCMを確実に検出できるようにするものである。

【0012】

【課題を解決するための手段】請求項1に記載の情報処理装置は、テレビジョン放送に関する信号を用いて閾値を決定する閾値決定手段と、閾値を用いてテレビジョン放送の無音区間を判定する判定手段と、テレビジョン放送のシーンチェンジを検出するシーンチェンジ検出手段と、シーンチェンジ検出手段の検出結果、および判定手段の判定結果に基づいてコマーシャルを検出するコマーシャル検出手段とを備えることを特徴とする。

【0013】請求項3に記載の情報処理方法は、テレビジョン放送に関する信号を用いて閾値を決定する閾値決定ステップと、閾値を用いてテレビジョン放送の無音区間を判定する判定ステップと、テレビジョン放送のシーンチェンジを検出するシーンチェンジ検出ステップと、シーンチェンジ検出ステップの検出結果、および判定ステップの判定結果に基づいてコマーシャルを検出するコマーシャル検出ステップとを含むことを特徴とする。

【0014】請求項4に記載の提供媒体は、テレビジョン放送に関する信号を用いて閾値を決定する閾値決定ステップと、閾値を用いてテレビジョン放送の無音区間を判定する判定ステップと、テレビジョン放送のシーンチェンジを検出するシーンチェンジ検出ステップと、シーンチェンジ検出ステップの検出結果、および判定ステップの判定結果に基づいてコマーシャルを検出するコマーシャル検出ステップとを含む処理を情報処理装置に実行させるコンピュータが読み取り可能なプログラムを提供することを特徴とする。

【0015】請求項1に記載の情報処理装置、請求項3に記載の情報処理方法、および請求項4に記載の提供媒体においては、テレビジョン放送に関する信号を用いて閾値が決定され、決定された閾値を用いてテレビジョン放送の無音区間が判定され、テレビジョン放送のシーンチェンジを検出されて、検出結果、および判定結果に基づいてコマーシャルが検出される。

【0016】

【発明の実施の形態】以下に本発明の実施の形態を説明するが、特許請求の範囲に記載の発明の各手段と以下の実施の形態との対応関係を明らかにするために、各手段の後の括弧内に、対応する実施の形態（但し一例）を付加して本発明の特徴を記述すると、次のようになる。但し勿論この記載は、各手段を記載したものに限定することを意味するものではない。

【0017】請求項1に記載の情報処理装置は、テレビ

ジョン放送に関する信号を用いて閾値を決定する閾値決定手段（例えば、図2の無音閾値決定部）と、閾値を用いて前記テレビジョン放送の無音区間を判定する判定手段（例えば、図2の無音区間検出部13）と、テレビジョン放送のシーンチェンジを検出するシーンチェンジ検出手段（例えば、図2のシーンチェンジ検出部15）と、シーンチェンジ検出手段の検出結果、および判定手段の判定結果に基づいてコマーシャルを検出するコマーシャル検出手段（例えば、図2のCM候補区間検出部16）とを備えることを特徴とする。

【0018】請求項2に記載の情報処理装置は、所定の区間のシーンチェンジの回数をカウントするカウント手段（例えば、図2のCM候補区間検出部16）をさらに備え、コマーシャル検出手段は、シーンチェンジの回数に基づいてコマーシャルを検出することを特徴とする。

【0019】本発明を適用したVCRの第1の構成例について、図1を参照して説明する。記録系のチューナ2は、端子1から入力されるテレビジョン放送のRF信号を復調し、得られた映像信号、音声信号、AGC信号、および音声多重モードを示す信号（以下、音声多重モード信号と記述する）を、CM検出回路3に出力する。また、チューナ2は、映像信号および音声信号を遅延回路4に出力する。

【0020】CM検出回路3は、チューナ2から入力された信号がCMであるか否かを判定し、その結果に対応し、スイッチ6へ制御信号として0または1を出力する。スイッチ6は、制御信号が1である場合、オフとされ、制御信号が0である場合、オンとされる。

【0021】遅延回路4は、CM検出回路3の処理に要する時間（本実施の形態においては、1分間）だけ、チューナ2から入力された映像信号および音声信号を遅延して変調回路5に出力する。変調回路5は、遅延回路4から入力された映像信号および音声信号を、所定の方式（例えば、MPEG2方式）で圧縮符号化し、さらに、所定の変調方法（例えば、EFM変調）で変調し、得られた変調信号をスイッチ6を介して書き込み回路7に出力する。書き込み回路7は、入力された変調信号を磁気テープ8に記録する。

【0022】再生系の読み取り回路9は、ユーザから入力されるコマンドに対応して、磁気テープ8に記録されている信号を読み取って復調し、図示せぬモニタに供給する。

【0023】なお、映像信号および音声信号を記録する媒体は、磁気テープに限定されるものではなく、光ディスク、光磁気ディスク、ハードディスク、半導体メモリ等であってもかまわない。

【0024】図2は、図1のCM検出回路3の第1の構成例を示している。このCM検出回路3において、チューナ2から入力される音声多重モード信号は無音閾値決定部11およびCM候補区間検出部16に供給され、映

像信号は遅延部14およびシーンチェンジ検出部15の供給され、音声信号はA/D変換部12に供給され、ACC信号は無音閾値決定部11に供給される。

【0025】無音閾値決定部11は、音声多重モード信号、A/D変換部12でデジタル化された音声信号、ACC信号、またはシーンチェンジ検出部15から入力される信号のいずれかをを用いて無音区間検出処理に用いられる閾値を演算し、無音区間検出部13に供給する。

【0026】無音区間検出部13は、A/D変換部12から入力されるデジタル化された音声信号のレベルと、無音閾値決定部11から供給された閾値を比較することにより、無音区間を検出し、その結果をシーンチェンジ検出部15に出力する。

【0027】シーンチェンジ検出部15は、同時に入力される2枚のフレーム画像（現フレームと、遅延部14により1フレーム時間（1/30秒）だけ遅延された前フレーム）を比較して、無音区間におけるシーンチェンジの有無を検出し、その結果をCM候補区間検出部16に出力する。

【0028】CM候補区間検出部16は、音声多重モード信号と、シーンチェンジ検出部15から供給されたシーンチェンジの情報をフレーム毎に2値化して、内蔵するメモリに記憶し、その情報に基づいてCM候補区間を検出し、CM候補区間においてはスイッチ6への制御信号として1を出力し、CM候補区間以外においては制御信号として0を出力する。なお、CM候補区間検出部16が内蔵するメモリには、過去1分間の情報を記憶させておく、その場合、内蔵されるメモリの容量は、60（秒）×30（フレーム）×2（データ）×1（ビット）となる。

【0029】次に、CM検出回路3の動作について、図3のフローチャートを参照して説明する。このCM検出処理は、チューナ2からCM検出回路3に各信号が入力されたときに開始される。ステップS1において、CM検出回路3の無音閾値決定部11は、予め演算した（後述）、無音区間検出処理（ステップS2）に用いられる閾値を無音区間検出部13に出力する。

【0030】ここで、無音閾値決定処理の詳細について、図4のフローチャートを参照して説明する。ステップS11において、A/D変換部12は、チューナ2から入力された所定の微小時間の音声信号（アナログ）を、所定のサンプリング周波数、所定の量子化レベルでデジタル化し、得られたデジタル音声信号（例えば、図5(A)）を無音閾値決定部11に出力する。無音閾値決定部11は、ステップS12において、A/D変換部12から入力されたデジタル音声信号を、図5(B)に示すように絶対値化し、ステップS13において、絶対値化されたサンプルのレベルの平均値（図5(C)）を演算する。

【0031】ステップS14において、無音閾値決定部

11は、ステップS13で得た平均値と、それまで記憶していた閾値とを比較して、小さい方の値を新たな閾値として記憶する。

【0032】ステップS15において、無音閾値決定部11は、全てのチャンネルに対してステップS11乃至S14の処理を実行したか否かを判定し、全てのチャンネルに対して処理を実行していないと判定した場合、ステップS16に進む。ステップS16において、無音閾値決定部11は、チューナ2にチャンネル切換信号を出力する。このチャンネル切換信号に対応して、チャンネルが切り替えられる。

【0033】その後、ステップS15において、全てのチャンネルに対して、ステップS11乃至S14の処理を実行したと判定された場合、無音閾値決定処理を終了する。なお、この無音閾値決定処理は、常に所定の間隔（例えば、10分毎）で繰り返して実行される。

【0034】なお、無音閾値決定処理の方法は、上述した方法以外にも考えられる。例えば、映像信号は存在するが、音声信号のレベルは確実に0となる、各チャンネルの放送開始時刻、および放送終了時刻において、音声信号を受信し、そのレベルnに所定のオフセット値Δを加えた値n+Δを無音閾値としてもよい。ただし、各チャンネルの放送開始時刻および放送終了時刻は既知であるものとする。

【0035】また、チューナ2からのACC信号を用いて無音閾値を決定してもよい。すなわち、チューナ2からのACC信号のレベルは、図6に示すように、ACC信号のレベルと音声信号のS/Nは反比例の関係にある。したがって、ACC信号のレベルから音声信号のS/Nを推定すれば、推定されたS/Nに対応する閾値を一意的に決定することができる。

【0036】さらに、音声信号のバズ成分は、映像信号の輝度レベルに依存しているので、映像信号の輝度レベルの平均値を用いて無音閾値を決定してもよい。例えば、輝度レベルの平均値が高い場合、音声信号のバズ成分が増加し、結果的に一定期間の音声信号のレベルが高くなるので、閾値を通常よりも高めに設定し、輝度レベルの平均値が低い場合、音声信号のバズ成分が減少し、結果的に一定期間の音声信号のレベルが低くなるので、閾値を通常よりも低めに設定する。なお、映像信号の輝度の平均値 $Y_A$ は、次式のように、シーンチェンジ検出部15で演算する。

$$Y_A = (\sum D_{ij}) / n \times m$$

ただし、 $i = 1$ 乃至 $n$ 、 $j = 1$ 乃至 $m$ であり、 $D_{ij}$ は、図7に示すように、音声信号に対応する画像の座標（ $i, j$ ）画素値である。

【0037】さらに、音声多重モード信号に基づいて無音閾値を決定してもよい。すなわち、音声多重モードが2カ国語放送からステレオに切り替わるときの音声信号のレベルを無音閾値としてもよい。

【0038】図3に戻る。ステップS2において、無音区間検出部13は、ステップS1で無音閾値決定部11から入力された閾値に基づいて無音区間を検出する。この無音区間検出処理の詳細について、図8のフローチャートを参照して説明する。

【0039】ステップS21において、A/D変換部12は、チューナ2から入力された、所定の微小時間の音声信号（アナログ）を所定のサンプリング周波数、および所定の量子化レベルでデジタル化し、得られたデジタル音声信号（例えば、図5(A))を無音閾値決定部11に出力する。無音区間検出部13は、ステップS22において、A/D変換部12から入力されたデジタル音声信号を、図5(B)に示すように絶対値化し、ステップS23において、絶対値化されたサンプルのレベルの平均値（図5(C))を演算する。

【0040】ステップS24において、無音閾値決定部11は、ステップS23で得た平均値が、無音閾値決定部11から入力された閾値よりも小さいか否かを判定し、平均値が閾値よりも小さいと判定した場合、ステップS25に進む。ステップS25において、無音区間検出部13は、この区間を無音区間と判定し、その情報をシーンチェンジ検出部15に出力する。

【0041】反対に、ステップS24において、平均値が閾値よりも小さくないと判定した場合、ステップS26に進む。ステップS26において、無音区間検出部13は、この区間を無音区間ではない（有音区間である）と判定し、その情報をシーンチェンジ検出部15に出力する。

【0042】図3のステップS3にリターンする。ステップS3において、シーンチェンジ検出部15は、入力された前後する2枚のフレームにおけるシーンチェンジを検出し、その結果をCM候補区間検出部16に出力する。このシーンチェンジ検出処理の詳細について、図9のフローチャートを参照して説明する。

【0043】ステップS31において、シーンチェンジ検出部15は、無音区間検出部13から入力された情報が無音区間を示すものであるか否かを判定し、入力された情報が無音区間を示す情報であると判定した場合、ステップS32に進む。

【0044】ステップS32において、シーンチェンジ検出部15は、入力された前後する2枚のフレーム画像の相関値Eを演算する。具体的には、図10に示すように、遅延部14を介して入力された画像（遅延画像）と、遅延部14を介さずに入力された画像（スルー画像）の対応する画素の画素値の差の絶対値の総和が相関値として次式により演算される。

$$E = \sum |D_{ij} - S_{ij}|$$

ただし、 $D_{ij}$ 、 $S_{ij}$ は、それぞれ、遅延画像またはスルー画像の座標（ $i$ 、 $j$ ）の画素値を表している。なお、この相関値は、フレーム間の相関の程度が小さいくなる

と、その値は大きくなり、フレーム間の相関の程度が大きくなると、その値は小さくなる。

【0045】ところで、相関値Eを演算する他の方法として、遅延画像とスルー画像の画素値のヒストグラムを用いる方法や、遅延画像とスルー画像を所定のブロックに分割し、ブロック毎に相関値を画素値の差分を求める方法を用いてもよい。

【0046】ステップS33において、シーンチェンジ検出部15は、ステップS32で得られた相関値が所定の閾値よりも大きいか否かを判定し、相関値が所定の閾値よりも大きい（フレーム間の相関の程度が小さい）と判定した場合、ステップS34に進む。

【0047】ステップS34において、シーンチェンジ検出部15は、入力された前後する2枚のフレーム間にはシーンチェンジが存在すると判定して、その情報をCM候補区間検出部16に出力する。

【0048】反対に、ステップS33において、相関値が所定の閾値よりも大きくない（フレーム間の相関の程度が大きい）と判定された場合、ステップS35に進む。

【0049】ステップS35において、シーンチェンジ検出部15は、入力された前後する2枚のフレーム間にはシーンチェンジがないと判定して、その情報をCM候補区間検出部16に出力する。

【0050】なお、ステップS31において、無音区間を示す情報ではないと判定された場合、その情報がCM候補区間検出部16に出力されて、図3のステップS4にリターンする。

【0051】ステップS4において、CM候補区間検出部16は、内蔵するメモリに2値化して記憶している過去1分間のフレーム単位の音声多重モード信号およびシーンチェンジ検出部15からの情報に基づいてCM候補区間を判定する。

【0052】すなわち、CM候補区間検出部16に内蔵されるメモリには、図11に示すように、音声多重モード信号（Audio\_Multi[]）については、ステレオモードが1、モノラルモードおよび2カ国語モードが0として記録され、シーンチェンジ検出部15から入力された情報（Scene\_Change[]）については、シーンチェンジであるフレームが1、シーンチェンジではないフレームが0として記録されている。

【0053】CM候補区間検出部16は、内蔵するメモリを参照し、図12(A)に示すように、シーンチェンジを示す信号が1であるフレーム（シーンチェンジ点）毎に区間（いまの場合、シーンチェンジ区間0乃至10）を区切り、各シーンチェンジ区間を構成するフレーム数を30で除算することにより、その区間の時間を演算する。また、CM候補区間検出部16は、図12(B)に示すように、音声多重モード信号が1である連続するフレームをステレオ区間とする。さらに、CM候補区間検出

部16は、図12(C)に示すように、単独のシーンチェンジ区間の時間、または隣接する複数のシーンチェンジ区間の合計時間が15秒の整数倍であり、且つ、ステレオ区間である区間をCM候補区間と判定し、CM候補区間においてはスイッチ6への制御信号として1を出力し、CM候補区間以外においてはスイッチ6への制御信号として0を出力する。

【0054】なお、スイッチ6は、この制御信号が1であるときオフとされ、0であるときオンとされるが、スイッチ6には、遅延回路4で1分間だけ遅延されたことにより、CM候補区間検出部16からの制御信号に同期している、変調された映像信号および音声信号が変調回路5から入力されているので、スイッチ6の後段には、本編の映像信号および音声信号だけが供給され、CM候補区間の映像信号および音声信号は供給されない。したがって、磁気テープ8には、テレビジョン放送のうちのCM候補区間を除いた本編だけが記録される。

【0055】次に、CM候補区間検出部16の他の動作例について説明する。この動作例は、CMが有する絶対的な特徴ではないが、CMの一般的な傾向（CMにおいてはシーンチェンジが複数回行われる）をCM候補区間の検出条件に追加したものである。

【0056】この動作例において、CM候補区間検出部16が内蔵するメモリには、図13に示すように、図11に示した情報に、無音区間検出部13の出力である無音区間を示す情報（No\_Sound[]）（有音部分は0、無音区間は1）が追加されて記録される。なお、この場合、内蔵されるメモリの容量は、

$60 \text{ (秒)} \times 30 \text{ (フレーム)} \times 3 \text{ (データ)} \times 1 \text{ (ビット)}$

だけ必要となる。

【0057】CM候補区間検出部16は、内蔵するメモリを参照し、図14(A)に示すように、無音区間を抽出し、図14(B)に示すように、シーンチェンジ点を抽出して、図14(C)に示すように、無音区間におけるシーンチェンジ点毎にシーンチェンジ区間（図14(C)の場合、シーンチェンジ区間0乃至10）を区切り、各シーンチェンジ区間を構成するフレーム数を30で除算することにより、その区間の時間を演算する。

【0058】また、CM候補区間検出部16は、図14(D)に示すように、単独のシーンチェンジ区間の時間、または隣接する複数のシーンチェンジ区間の合計時間が15秒の整数倍となるようにグループ化し、図14(E)に示すように、音声多重モード信号が1である区間を抽出し、図14(F)に示すように、15秒の整数倍にグループ化され、且つ、ステレオ区間である区間（いまの場合、区間a乃至c、および区間d、e）をCM候補区間と判定する。さらに、CM候補区間検出部16は、図14(G)に示すように、隣接するCM候補区間の両端のCM候補区間（いまの場合、CM候補区間a、c、d、

e）におけるシーンチェンジの回数を所定の閾値（例えば、1）と比較して、シーンチェンジの回数が閾値よりも大きいCM候補区間（いまの場合、CM候補区間a、c）をCM区間と判定する。なお、CM候補区間a、cに挟まれているCM候補区間bもCM区間と判定される。

【0059】CM候補区間検出部16は、CM区間においてはスイッチ6への制御信号として1を出力し、CM区間ではない区間においてはスイッチ6への制御信号として0を出力する。

【0060】次に、CM候補区間検出部16のさらに他の動作例について説明する。この動作例は、CMが有する絶対的な特徴が変更された場合（例えば、CMの時間が15秒の整数倍から、14秒の整数倍に変更された等）、特徴の変更に対応してCM候補区間の判定に用いる基準値を変更するものである。

【0061】この動作例において、CM候補区間検出部16は、内蔵するメモリ（図11）を参照し、図15(A)に示すように、シーンチェンジを示す信号が1であるフレーム毎にシーンチェンジ区間（いまの場合、シーンチェンジ区間0乃至10）を区切り、各シーンチェンジ区間を構成するフレーム数を30で除算することにより、その区間の時間を演算する。また、CM候補区間検出部16は、図15(B)に示すように、音声多重モード信号が1である連続するフレームをステレオ区間とする。

【0062】ここで、CM候補区間検出部16は、ステレオ区間の時間と、そのステレオ区間に隣接するステレオ区間ではない区間の時間を比較し、ステレオ区間の時間が十分に短い場合、そのステレオ区間（例えば、シーンチェンジ区間1乃至4、シーンチェンジ区間8、9）をCM候補区間と判定する。

【0063】次に、CM候補区間検出部16は、CM候補区間（シーンチェンジ区間1乃至4）の合計時間が56秒であることから、CMの時間は7秒単位、14秒単位、または28秒単位のいずれかであると判定し、さらに、CM候補区間検出部16は、シーンチェンジが14秒または28秒の間隔で行われていることから、最終的に、CMの時間が14秒の整数倍であると判定する。

【0064】以上のような結果（CMの時間が14秒の整数倍であると判定される）が所定の回数以上連続して得られた場合、CM候補区間検出部16は、CMの時間は14秒の整数倍に変更されたと判定して、以後、14秒の整数倍をCM区間検出の基準値として使用する。

【0065】なお、CMが有する絶対的な特徴が変更された場合（上述したようなCM時間の変更や欧米におけるブラックフレーム、ブルーフレーム挿入の廃止等）、変更された絶対的な特徴に対応して変更したCM区間判定用の基準値を、装置外部からCM候補区間検出部16に供給するようにしてもよい。この場合、CM候補区間

検出部16においては、CM区間検出用の基準値をフラッシュメモリのような書き換え可能な記録媒体に格納しておけばよい。

【0066】次に、CM検出回路3の第2の構成例について、図16を参照して説明する。この構成例は、図2に示した構成例にCMブロック検出部21を追加したものである。CMブロック検出部21は、CM候補区間検出部16で検出された単独のCM候補区間のフレーム数を用いて、本編と本編の間に挟まれた複数のCMからなるCMブロックを検出し、その結果に基づいてスイッチ6へ制御信号を出力する。

【0067】この構成例によれば、以下に記述するようなCMの特徴を利用してCMブロックを検出することができる。すなわち、複数のCMが連続して放送される場合、個々のCMとしては標準的なフレーム数に対して5フレーム程度の誤差を有しているが、CMブロック（連続する複数のCM）としては、標準的なフレーム数に対する誤差が3フレーム程度である。例えば、15秒間のCMが4本連続して放送される場合、個々のCMのフレーム数は、 $450 \pm 5 (= 15 \times 30 \pm 5)$  であるが、CMブロックのフレーム数は、 $1800 \pm 20 (= 15 \times 4 \times 30 \pm 5 \times 4)$  とはならず、 $1800 \pm 3$  となる。

【0068】次に、その動作について図17、図18を参照して説明する。CMブロック検出部21は、CM候補区間検出部16が検出した各CM候補区間（図17(A)のシーンチェンジ区間1+2, 3, 4, 8, 9)について、そのフレーム数をCM候補区間検出部16から読み出し、隣接するCM候補区間のフレーム数を合計して、その値が上述した誤差の許容範囲（ $\pm 3$ ）に収まっているか否かを判定する。CMブロック検出部21は、合計したフレーム数が誤差の許容範囲に収まっていると判定した場合、その隣接するCM候補区間をCMブロックであると判定し、反対に、合計したフレーム数が誤差の許容範囲に収まっていないと判定した場合、その隣接するCM候補区間をCMブロックではないと判定する。

【0069】いまの場合、図17(D)に示すように、隣接するCM候補区間（シーンチェンジ区間1乃至4）のフレーム数の合計は、 $1803 (= 453 + 446 + 904)$  であり、標準的なフレーム数1800（ $= 60 \text{ 秒} \times 30 \text{ フレーム}$ ）に対する誤差（3）は、許容範囲内であるので、隣接するCM候補区間（シーンチェンジ区間1乃至4）は、CMブロックであると判定される。

【0070】一方、隣接するCM候補区間（シーンチェンジ区間8, 9）のフレーム数の合計は、 $910 (= 455 + 455)$  であり、標準的なフレーム数900（ $= 30 \text{ 秒} \times 30 \text{ フレーム}$ ）に対する誤差（10）は、許容範囲外であるので、隣接するCM候補区間（シーンチェンジ区間8, 9）は、CMブロックではないと判定される。

【0071】図18(C)に示すようなCM候補区間がCM候補区間検出部16から入力された場合、CMブロック検出部21は、上述した図17の場合と同様に、シーンチェンジ区間1乃至4（CMブロックAとする）、およびシーンチェンジ区間7乃至10（CMブロックBとする）を、CMブロックであると判定する。

【0072】さらに、CMブロック検出部21は、CMブロックA, Bに挟まれたCM候補区間ではない区間（シーンチェンジ区間5, 6）のフレーム数をCM候補区間検出部16から読み出して、そのフレーム数（いまの場合、150）が所定の閾値（例えば、300フレーム（10秒））以下である場合、その区間は本編ではなく、ユーザにとってはCMと同様なものであると判断して、前後するCMブロックA, Bに包含する。すなわち、CMブロック検出部21は、図18(D)に示すように、シーンチェンジ区間1乃至10を、1つのCMブロックであると判定する。

【0073】なお、CMブロック検出部21は、CMブロックにおいてはスイッチ6への制御信号として1を出力し、CMブロック以外においてはスイッチ6への制御信号として0を出力する。

【0074】次に、CM検出回路3の第3の構成例について、図19を参照して説明する。この構成例は、図2に示した第1の構成例にCM特徴量検出部31を追加したものである。CM特徴量検出部31には、前後する2枚のフレーム画像（映像信号）、デジタル化された音声信号、CM候補区間検出部16に内蔵されるメモリに記憶されている情報（図11）、およびCM候補区間検出部16が検出したCM候補区間の情報が供給される。CM特徴量検出部31は、供給される情報が、CMの様々な特徴を有しているか否かを判定し、その結果に基づいて、CM区間を検出する。さらに、CM特徴量検出部31は、検出結果に対応してスイッチ6に制御信号を出力する。

【0075】CM特徴量検出部31の動作について、図20のフローチャートを参照して説明する。この特徴量検出処理は、CM候補区間検出部16が検出した各CM候補区間に対して実行される。

【0076】ステップS41において、CM特徴量検出部31は、評価値を0に初期化する。ステップS42において、CM特徴量検出部31は、CM候補区間検出部16から入力されたCM候補区間のシーンチェンジ回数をカウントして、シーンチェンジ回数が所定の閾値（15秒あたり5回）以上であるか否かを判定し、計測したシーンチェンジ回数が所定の閾値以上であると判定した場合、評価値に1を加算する。反対に、計測したシーンチェンジ回数が所定の閾値以上ではないと判定された場合、評価値への加算は行われない。この処理は、CMにおいては頻繁にシーンチェンジが行われるという、一般的な特徴に基づくものである。



【0077】ステップS42において、CM特徴量検出部31は、CM候補区間の音声信号の周期性を検出する。この音声信号の周期性検出処理は、CMにおいてはBGMが用いられるという、一般的な特徴に基づくものである。

【0078】この音声信号の周期性検出処理について、図21のフローチャートを参照して説明する。ステップS61において、CM特徴量検出部31は、図22に示すように、A/D変換部12から入力された音声信号のレベルが所定の閾値以上である点とピークとして検出する（図22(B)）。ただし、図22(A)は音声信号をアナログで表示している。

【0079】ステップS62において、CM特徴量検出部31は、ステップS61で検出したピークが所定の期間（数秒）よりも長く周期性を有しているか否かを判定する。周期性を判定するには、例えば、FFT(Fast Fourier Transform)を用いてもよいし、単にピークの間隔を計測してもよい。検出したピークが所定の期間よりも長く周期性を有していると判定された場合、ステップS63に進む。

【0080】ステップS63において、CM特徴量検出部31は、評価値に1を加算する。

【0081】ステップS62において、検出したピークが所定の期間よりも長く周期性を有していないと判定された場合、評価値への加算は行われないので、ステップS63はスキップされる。

【0082】図20のステップS44にリターンする。ステップS44において、CM特徴量検出部31は、CM候補区間の音声信号および映像信号の連続性を検出する。この連続性検出処理について、図23のフローチャートを参照して説明する。

【0083】ステップS71において、CM特徴量検出部31は、音声信号のピークの周期性が、CM候補区間の95%以上の時間よりも長く継続しているか否かを判定する。ピークの周期性の継続時間の計測については、上述したステップS43で得られた情報を利用する。音声信号のピークの周期性が、CM候補区間の95%以上の時間よりも長く継続していると判定された場合、ステップS72に進む。

【0084】ステップS72において、CM特徴量検出部31は、評価値に1を加算する。

【0085】ステップS73において、CM特徴量検出部31は、前後する2枚の画像の対応する画素の画素値の差分を演算し、その差分が所定に閾値以下である画素、すなわち、2枚の画像間で動きがない画素を検出する。ステップS74において、CM特徴量検出部31は、ステップS73で検出した画素の画素値のヒストグラムを生成し、その最大値を示す画素値を背景の代表色として検出する。なお、ステップS73、S74の処理は、CM候補区間において、所定の間隔で繰り返し実行

される。

【0086】ステップS75において、CM特徴量検出部31は、ステップS74で検出した複数の背景の代表色を参照し、同じ画素値が連続して背景の代表色として検出されているか否かを判定する。同じ画素値が連続して背景の代表色として検出されていると判定された場合、ステップS76に進む。

【0087】ステップS76において、CM特徴量検出部31は、評価値に1を加算する。

【0088】なお、ステップS71において、音声信号のピークの周期性が、CM候補区間の95%以上の時間よりも長く継続していないと判定された場合、評価値に加算を行わないので、ステップS72はスキップされる。

【0089】また、ステップS75において、同じ画素値が連続して背景の代表色として検出されていないと判定された場合、評価値への加算は行われないので、ステップS76はスキップされる。

【0090】図20のステップS45にリターンする。ステップS45において、CM特徴量検出部31は、画像の再現性を検出する。この再現性検出処理は、例えば、15秒間のケチャップのCMが、緑色の森のシーン（3秒）、青色の空のシーン（2秒）、緑色の森のシーン（2秒）、赤色のケチャップのシーン（3秒）、緑色の森のシーン（2秒）、および、赤色のケチャップのシーン（3秒）で構成されるように、CMにおいては、同じ画像が再現されるという、一般的な特徴に基づくものである。

【0091】この画像の再現性検出処理について、図24のフローチャートを参照して説明する。ステップS81において、CM特徴量検出部31は、CM候補区間をシーンチェンジ点で区分する。ステップS82において、CM特徴量検出部31は、ステップS81で区分した各シーンの画素値のヒストグラムを16階調で生成し、最大値を示す画素値を、そのシーンの代表色として検出する。

【0092】ステップS83において、CM特徴量検出部31は、CM候補区間において、シーンの代表色が再現されているか（異なるシーンで同じ代表色が検出されているか）否かを判定する。シーンの代表色が再現されていると判定された場合、ステップS84に進む。

【0093】ステップS84において、CM特徴量検出部31は、評価値に1を加算する。

【0094】なお、ステップS83において、シーンの代表色が再現されていないと判定された場合、評価値への加算は行われないので、ステップS84はスキップされる。

【0095】図20のステップS46にリターンする。ステップS46において、CM特徴量検出部31は、図26(A)に示すようなテロップ（字幕）を検出する。こ

のテロップ検出処理は、CMの終了時にはテロップが表示するという、一般的な特徴に基づくものである。

【0096】このテロップ検出処理について、図25のフローチャートを参照して説明する。ステップS91において、CM特徴量検出部31は、CM候補区間の終了側の所定の時間（例えば、5秒間）の各フレーム画像において、図26(B)に示すように、等しい画素値を持ち、且つ、所定の数よりも多く隣接している画素を検出する。

【0097】ステップS92において、CM特徴量検出部31は、図26(C)に示すように、ステップS91で検出した画素と等しい画素値をもつ画素を検索して領域化する。

【0098】ステップS93において、CM特徴量検出部31は、ステップS92で得られた領域が時間方向に連続しているか（前後するフレームにも同じ領域が存在するか）否かを判定する。領域が時間方向に連続していると判定された場合、ステップS94に進む。

【0099】ステップS94において、CM特徴量検出部31は、画像上の全ての領域の重心を検出し、その重心と領域上の各点との距離の平均値を演算して、さらに、その平均値と所定の閾値とを比較することにより、領域が密集しているか否かを判定する。領域が密集していると判定された場合、ステップS95に進む。

【0100】ステップS95において、CM特徴量検出部31は、評価値に1を加算する。

【0101】なお、ステップS93において、領域が時間方向に連続していないと判定された場合、評価値への加算は行われないので、ステップS95はスキップされる。

【0102】図20のステップS47に戻る。ステップS47において、CM特徴量検出部31は、上述したテロップ検出処理と同様に方法を用いて、CM候補区間の開始から終了までの全区間において、画像上に文字（テロップを含む）が所定の時間よりも長く存在するか否かを判定する。画像上に文字が所定の時間よりも長く存在すると判定された場合、評価値に1を加算される。反対に、画像上に文字が所定の時間よりも長く存在しないと判定された場合、評価値への加算は行われず、この処理は、CMにおいては文字が表示されるという、一般的な特徴に基づくものである。

【0103】ステップS48において、CM特徴量検出部31は、CM候補区間の開始点および終了点以外の無音区間の割合を検出する。この無音率検出処理は、CMにおいては無音区間が少ないという、一般的な特徴に基づくものである。無音率検出処理について、図27のフローチャートを参照して説明する。

【0104】ステップS101において、CM特徴量検出部31は、CM候補区間検出部16に内蔵されているメモリに記憶されている情報（図11）を読み出して、

図28に示すように、CM候補区間の開始点および終了点以外の無音区間のフレーム数を検出する。さらに、CM特徴量検出部31は、江検出した無音区間のフレーム数を、CM候補区間を構成する全フレーム数で除算することにより、無音率を演算する。

【0105】ステップS102において、CM特徴量検出部31は、ステップS101で演算した無音率が所定の閾値よりも小さいか否かを判定する。無音率が所定の閾値よりも小さいと判定された場合、ステップS103に進む。

【0106】ステップS103において、CM特徴量検出部31は、評価値に1を加算する。

【0107】なお、ステップS102において、無音率が所定の閾値よりも小さくないと判定された場合、評価値への加算は行われないので、ステップS103はスキップされる。

【0108】図20のステップS49にリターンする。ステップS49において、CM特徴量検出部31は、評価値が所定の閾値よりも大きいと判定し、評価値が所定の閾値よりも大きいと判定した場合、ステップS50に進む。ステップS50において、CM特徴量検出部31は、CM候補区間をCM区間であると判定する。

【0109】なお、ステップS49において、評価値が所定の閾値よりも大きくないと判定された場合、ステップS51に進む。ステップS51において、CM特徴量検出部31は、CM候補区間をCM区間ではないと判定する。

【0110】この判定結果に対応して、CM特徴量検出部31は、CM区間においてはスイッチ6への制御信号として1を出力し、CM区間以外においてはスイッチ6への制御信号として0を出力する。

【0111】以上のように、本実施の形態であるVCR（図1）においては、テレビジョン放送のうちの本編だけが磁気テープ8に録画され、CMは録画されない。したがって、この磁気テープ8を再生した場合、本編だけが連続して表示される。

【0112】ところで、CMを再生時には早送りするものの、録画はしておきたいというユーザのニーズが存在するので、そのようなニーズに対応する、本発明を適用したVCRの第2の構成例について、図29を参照して説明する。この構成例は、図1に示した構成例からスイッチ6を削除し、CM検出回路3のCM検出結果を書き込み回路7に供給するようにしたものである。

【0113】この構成例の記録系において、書き込み回路7は、変調回路5から入力される全て（テレビジョン放送の本編とCM）の映像信号と音声信号を磁気テープ8に記録とともに、CM検出回路3からの情報に基づいて、磁気テープ8の所定の位置に、CMの関する情報（CMの時間的な位置等）を記録する。

【0114】再生系において、読み取り回路9は、ユー

ザから本編だけを再生するコマンド（CMを除去するコマンド）が入力された場合、磁気テープ9の所定の位置に記録されているCMに関する情報に基づいてCMを除去（早送り等）し、本編だけを復調して図示せぬモニタに供給する。

【0115】なお、本発明の適用は、VCRに限定されるものではなく、テレビジョン受像機やチューナ等に適用することが可能である。

【0116】また、上記各処理を行うコンピュータプログラムは、磁気ディスク、CD-ROM等の情報記録媒体よりなる提供媒体のほか、インターネット、デジタル衛星などのネットワーク提供媒体を介してユーザに提供することができる。

【0117】

【発明の効果】以上のように、請求項1に記載の情報処理装置、請求項3に記載の情報処理方法、および請求項4に記載の提供媒体によれば、無音区間およびシーンチェンジに基づいてコマercialを検出するようにしたので、テレビジョン放送に含まれるCMを確実に検出することが可能となる。

【図面の簡単な説明】

【図1】本発明を適用したVCRの第1の構成例を示すブロック図である。

【図2】図1のCM検出回路3の第1の構成例を示すブロック図である。

【図3】CM検出回路3の動作を説明するフローチャートである。

【図4】図2の無音閾値決定部11の処理を説明するフローチャートである。

【図5】図2の無音閾値決定部11の動作を説明するための図である。

【図6】図2の無音閾値決定部11の動作を説明するための図である。

【図7】図2の無音閾値決定部11の動作を説明するための図である。

【図8】図3のステップS2の無音区間検出処理を説明するフローチャートである。

【図9】図3のステップS3のシーンチェンジ検出処理を説明するフローチャートである。

【図10】シーンチェンジ検出処理を説明するための図である。

【図11】CM候補区間検出部16に内蔵されるメモリに記録されている情報を説明するための図である。

【図12】CM候補区間検出部16の処理を説明するための図である。

【図13】CM候補区間検出部16に内蔵されるメモリに記録されている情報を説明するための図である。

【図14】CM候補区間検出部16の処理を説明するための図である。

【図15】CM候補区間検出部16の処理を説明するための図である。

【図16】図1のCM検出回路3の第2の構成例を示すブロック図である。

【図17】図16のCMブロック検出部21の処理を説明するための図である。

【図18】図16のCMブロック検出部21の処理を説明するための図である。

【図19】図1のCM検出回路3の第3の構成例を示すブロック図である。

【図20】図19のCM特徴量検出部31の動作を説明するフローチャートである。

【図21】図20のステップS43の音声信号の周期性検出処理を説明するフローチャートである。

【図22】音声信号の周期性検出処理を説明するための図である。

【図23】図20のステップS44の連続性検出処理を説明するフローチャートである。

【図24】図20のステップS45の再現性検出処理を説明するフローチャートである。

【図25】図20のステップS46のテロップ検出処理を説明するフローチャートである。

【図26】テロップ転出処理を説明するための図である。

【図27】図20のステップS48の無音率検出処理を説明するフローチャートである。

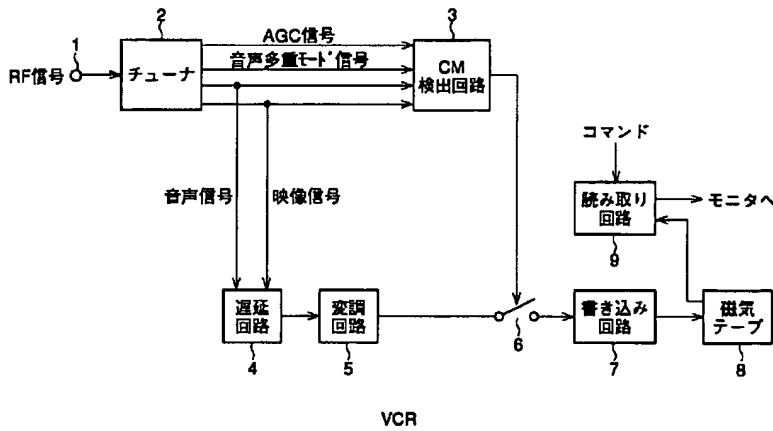
【図28】無音率検出処理を説明するための図である。

【図29】本発明を適用したVCRの第2の構成例を示すブロック図である。

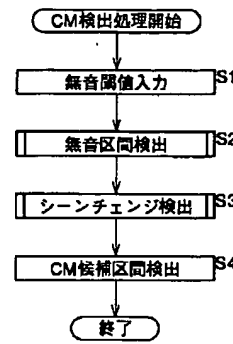
【符号の説明】

2 チューナ, 3 CM検出回路, 4 遅延回路, 5 変調回路, 6 スイッチ, 7 書き込み回路, 8 磁気テープ, 9 読み取り回路, 11 無音閾値決定部, 12 A/D変換部, 13 無音区間検出部, 14 遅延部, 15 シーンチェンジ検出部, 16 CM候補区間検出部, 21 CMブロック検出部, 31 CM特徴量検出部

【図1】



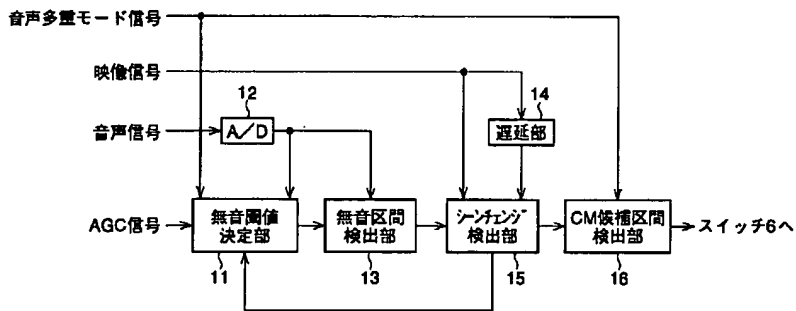
【図3】



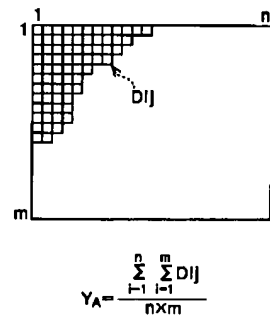
【図7】

遅延映像

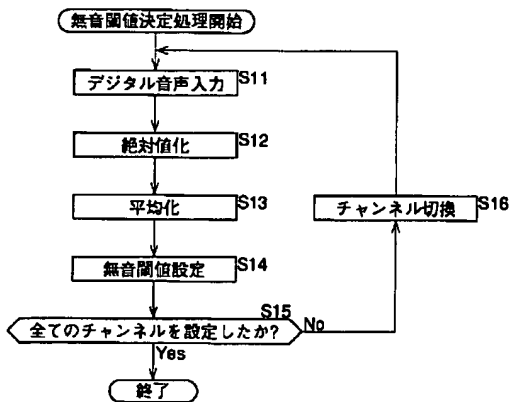
【図2】



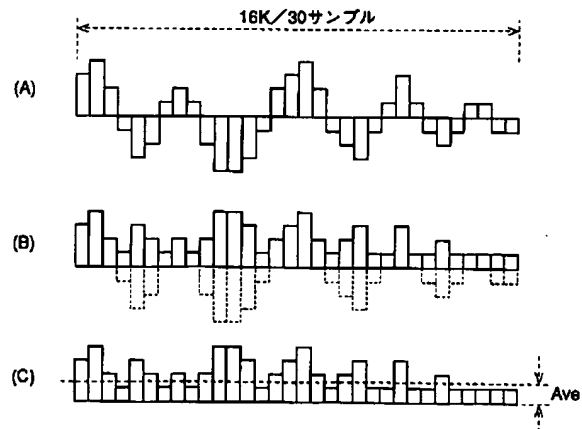
CM検出回路3



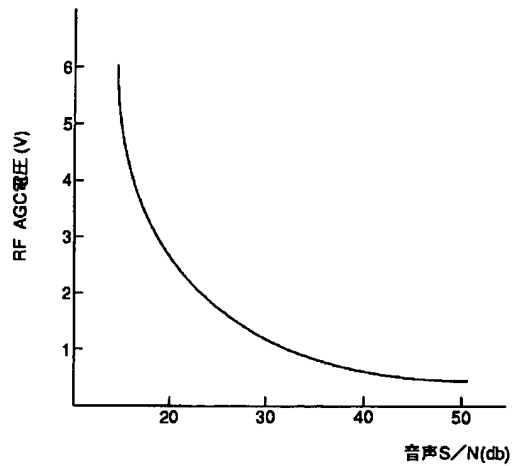
【図4】



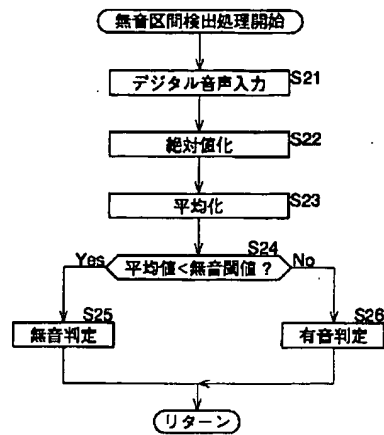
【図5】



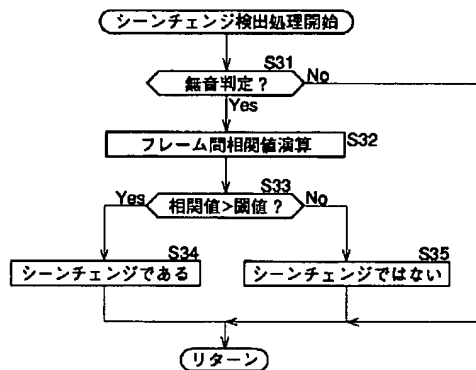
【図6】



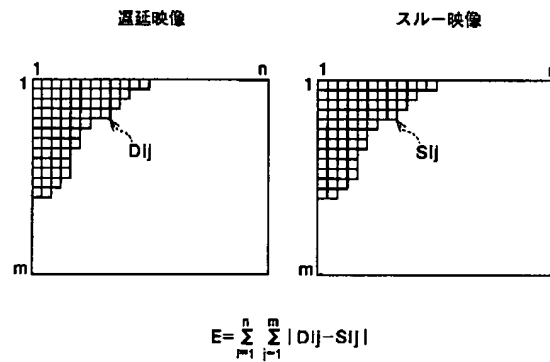
【図8】



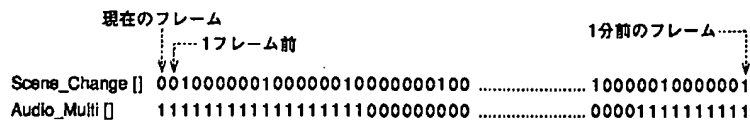
【図9】



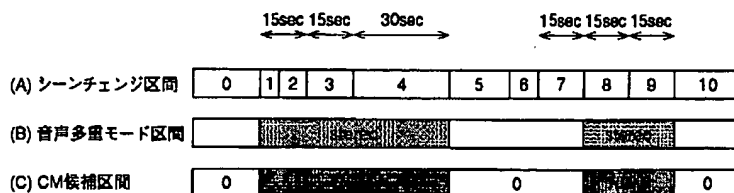
【図10】



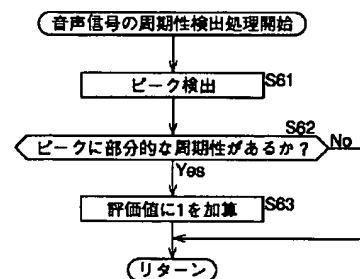
【図11】



【図12】



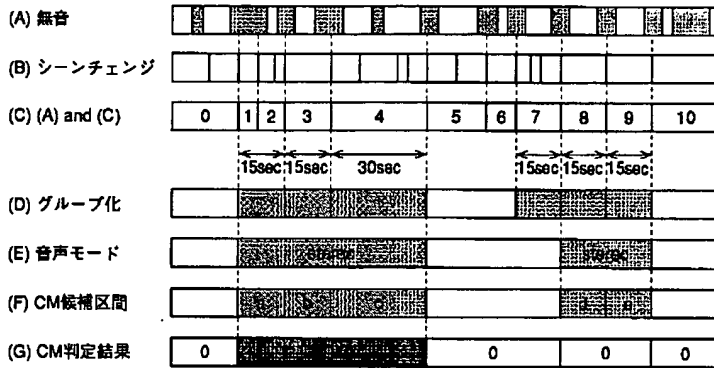
【図21】



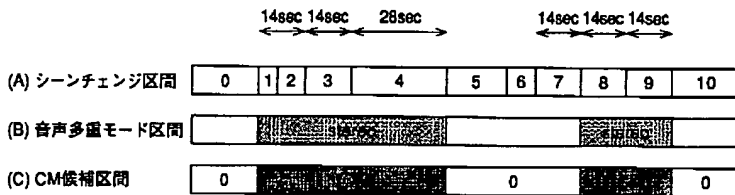
【図13】

	現在のフレーム	1分前のフレーム
No_Sound []	011110001110000111100001100 .....	00001111000011
Scene_Change []	0010000001000000010000000100 .....	10000010000001
Audio_Multi []	11111111111111111000000000 .....	00001111111111

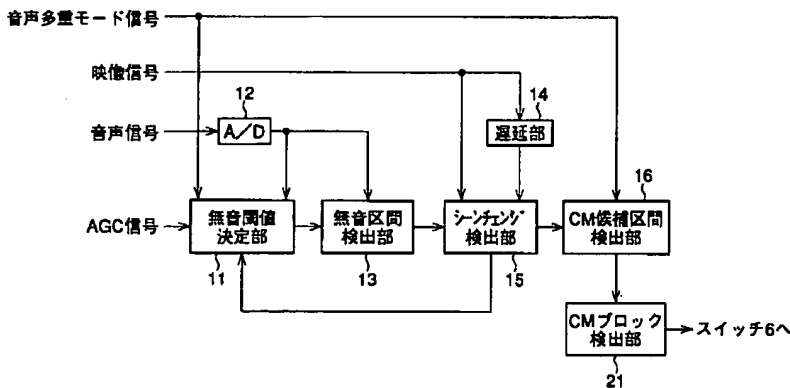
【図14】



【図15】

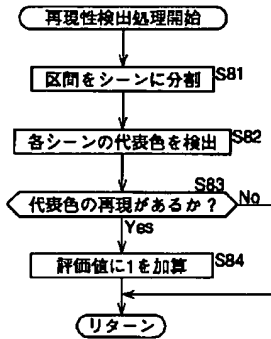


【図16】

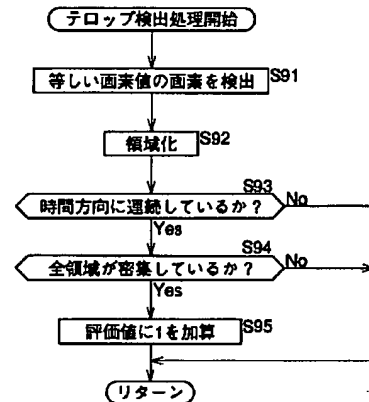


CM検出回路3

【図24】

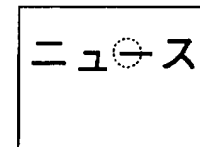


【図25】

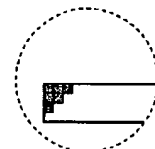


【図26】

(A)



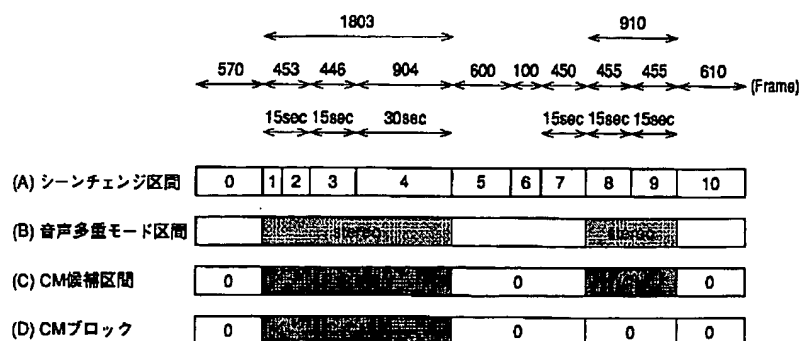
(B)



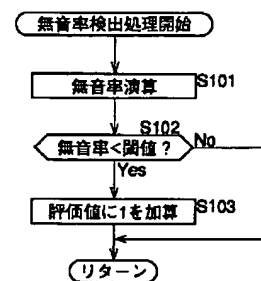
(C)



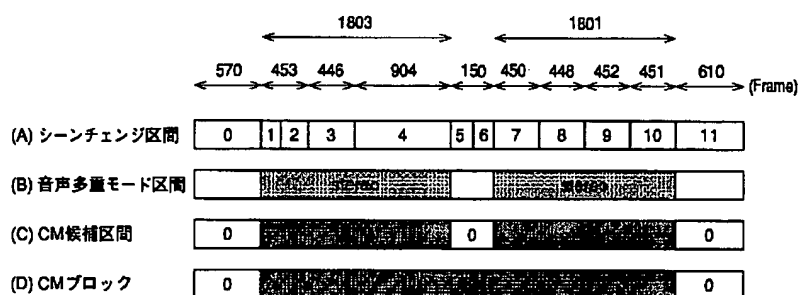
【図17】



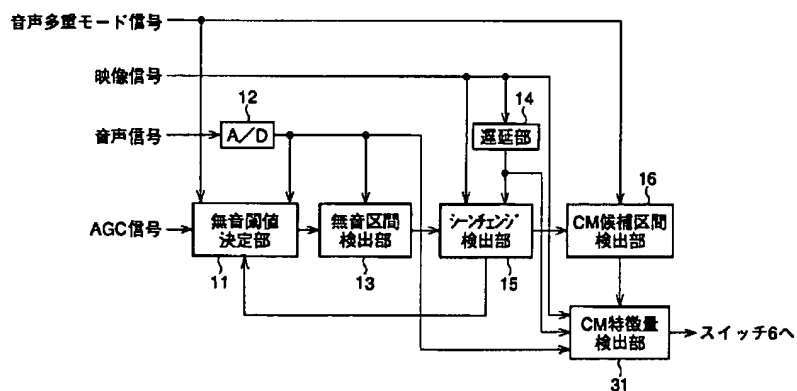
【図27】



【図18】

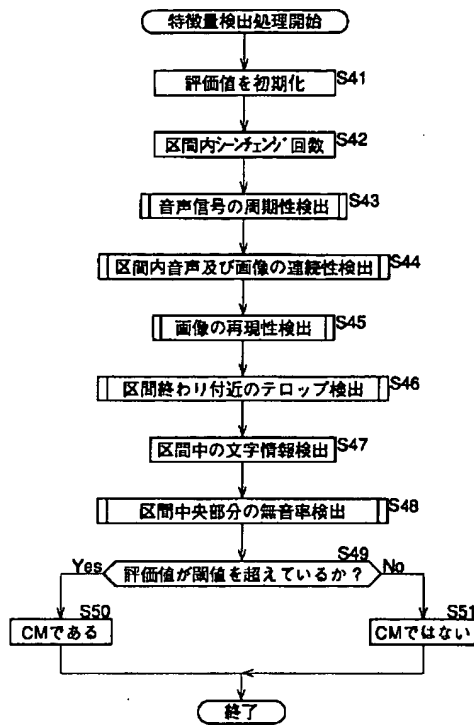


【図19】

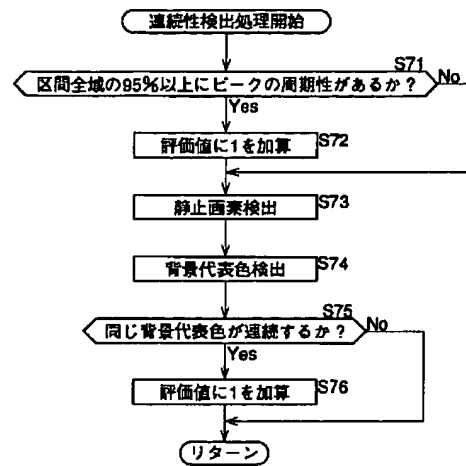


CM検出回路3

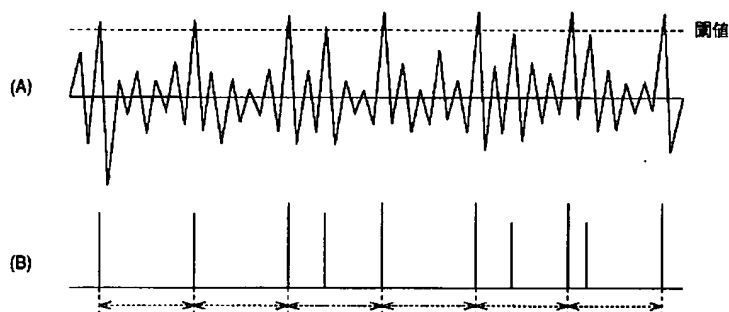
【図20】



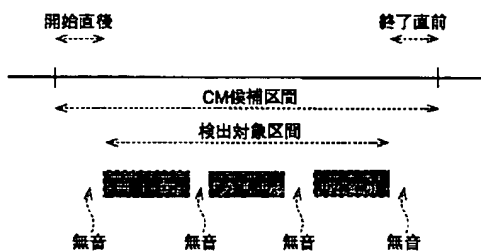
【図23】



【図22】

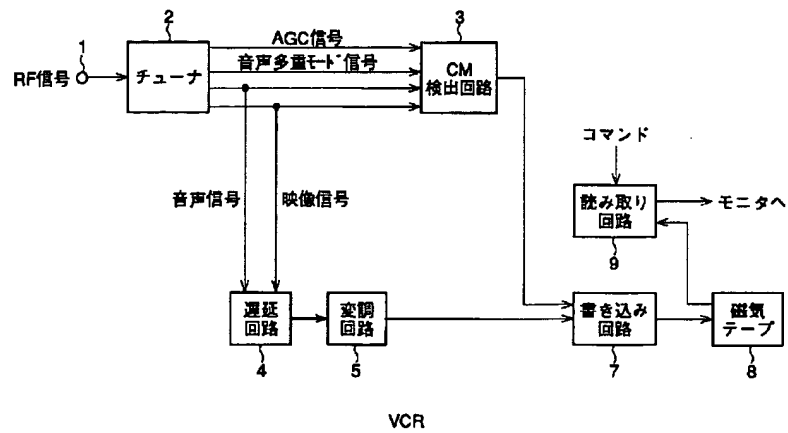


【図28】





【図29】



フロントページの続き

(72)発明者 宮田 勝成  
東京都品川区北品川6丁目7番35号 ソニ  
ー株式会社内

Fターム(参考) 5C018 AA04 AB09 AB11 AC00 HA01  
HA08